



CENTRE PERELMAN
DE PHILOSOPHIE DU DROIT

**De la donnée à la décision :
comment réguler par les données et
les algorithmes**

Par David Restrepo Amariles & Gregory
Lewkowicz

A paraître dans *Les big data à découvert*, Paris, CNRS Editions, 2016



De la donnée à la décision : comment réguler par les données et les algorithmes

Par David Restrepo Amariles & Gregory Lewkowicz

En droit, le développement des Big Data soulève au moins deux questions complémentaires. La première concerne la protection des données et de la vie privée . La deuxième, dont il sera question ici, porte sur les transformations du droit induites par le Big Data.

L'engouement pour les perspectives offertes par le Big Data dans le champ juridique s'inscrit dans un mouvement de fond, parfois qualifié de « tournant mathématique du droit ». Il vise à améliorer la qualité et l'efficacité du droit en mobilisant les ressources des mathématiques et des sciences de l'ingénieur. Il s'illustre par le recours croissant à des études quantitatives, par la mesure du droit et de son impact par des indicateurs ou encore par l'application automatique du droit par des algorithmes. Le Big Data complète, dans ce contexte, l'arsenal des autorités publiques et des praticiens du droit.

Le recours au Big Data en droit concerne surtout les domaines dans lesquels d'abondantes données structurées sont disponibles, tels que le droit bancaire et financier ou le droit de la sécurité sociale. Ce sont pourtant les initiatives fiscales qui ont valeur d'exemple par leur nombre. Une étude sur les pratiques fiscales de 86 États révélait ainsi que 69 d'entre eux avaient, en 2014, eu recours aux Big Data¹. L'analyse de la stratégie de l'autorité fiscale britannique (*Her Majesty's Revenue & Customs – HRMC*), fondée depuis 2009 sur l'exploitation d'une base de données et d'un logiciel d'analyse (un système dénommé « Connect »), permet ainsi de circonscrire les effets de la régulation par les données et les algorithmes. En France, la mise en place d'une stratégie similaire au moyen du *logiciel de ciblage de la fraude et valorisation des requêtes* en est encore à ses balbutiements.

Big Data et droit fiscal

Grâce au Big Data, le logiciel Connect fait émerger de l'analyse de très nombreuses données des modèles de comportement de fraude ou d'évasion fiscale et identifie les déclarations fiscales suspectes. Il permet au HRMC d'ouvrir des enquêtes au départ d'indices sérieux de fraude. Il modifie également la manière dont les agents du fisc surveillent le respect du droit fiscal.

Premièrement, le logiciel met à leur disposition des données sur les revenus et les modes de vie des contribuables auparavant difficilement accessibles et qui rendent pour partie obsolète l'examen manuel des dossiers. Connect utilise des techniques de *text* et *data mining* (cf. IV.7) afin d'extraire, d'une part, des informations contenues dans plus de 30 bases de données structurées, telles que celles des fiches de paie ou du registre foncier, et d'autre part, des données non-structurées, telles que celles issues des réseaux sociaux ou des sites de ventes aux enchères. Un contribuable qui rend public des photographies témoignant d'un mode de vie extravagant par rapport à ses revenus déclarés, ou qui met en location un bien à l'étranger qu'il ne déclare pas, est ainsi signalé par Connect comme une cible pour l'administration.

Deuxièmement, le logiciel limite le pouvoir d'appréciation de l'agent en charge du contrôle quant à l'opportunité d'enquêter. La confiance dans la puissance d'exploration des données de Connect rend en effet difficile de justifier d'écarter l'examen d'un dossier identifié par le logiciel. En 2014, 83 % des enquêtes fiscales ouvertes au Royaume-Uni ont d'ailleurs fait suite à un signalement du logiciel.

¹ *EY Global Survey :VAT/GST electronic filing and data extraction*, Ernst&Young, 2014, p. 36.

Troisièmement, Connect automatise l'application du droit et limite d'autant le pouvoir d'interprétation de la règle par l'agent du fisc. En effet, les algorithmes mis en œuvre intègrent une certaine interprétation des règles juridiques, en fonction de laquelle ils extraient et analysent les données. A terme, le logiciel a d'ailleurs vocation à établir le montant dont le contribuable devrait s'acquitter en appliquant directement les règles fiscales aux éléments factuels collectés. Pour les impôts indirects, on observe d'ailleurs en 2016 que les débats entre le fisc et les contribuables portent de moins en moins sur la soumission à l'impôt de telle ou telle transaction, mais sur le périmètre de l'ensemble de données le plus pertinent pour établir l'impôt.

De la règle au code

Le cas du programme Connect montre que l'utilisation du Big Data dans le domaine fiscal, de la collecte des données à la décision, n'affecte pas uniquement les méthodes de contrôle. En limitant le rôle de l'examen manuel, en définissant le périmètre des informations pertinentes, en identifiant les contribuables à contrôler et en appliquant une certaine interprétation du droit fiscal pour déterminer l'impôt, l'initiative du HRMC modifie le contenu de la règle fiscale parce qu'il en modifie fondamentalement l'application. En pratique, la règle est remplacée par les normes mises en œuvre par le logiciel et les indices de fraude qu'il utilise déterminent les standards à respecter. Les contribuables y adaptent d'ailleurs leur comportement ainsi que leur argumentation dans leurs litiges avec le fisc.

L'exemple fiscal met en exergue les implications générales de la régulation par les données et les algorithmes. Premièrement, la prépondérance de l'opinion experte cède le pas aux données. La rédaction d'un contrat respectant les standards de marché ne se fera plus sur la base de l'expertise de l'avocat chargé de sa rédaction, mais sur la base d'un logiciel tel que Bloomberg Law : Coporate Transactions qui vérifiera les variations entre les clauses du contrat proposé et celles qui sont statistiquement les plus fréquentes dans ses bases de données. Cette évolution n'est pas neutre sur le plan des valeurs véhiculées par le droit. On a ainsi pu montrer que dans le domaine des contrôles policiers, l'utilisation du Big Data pouvait renforcer les discriminations à l'égard de groupes défavorisés de la société en amplifiant celles dont les données de départ portent la trace. Deuxièmement, le droit cesse d'appréhender tout cas comme l'application contestable de règles aux faits de la cause. Le Big Data présuppose un monde déjà juridiquement chargé, où l'appréhension des faits, leur qualification et leur examen à la lumière des règles sont intriqués dans un même processus et quasiment indiscernables. Le code et les opérations du logiciel remplacent aussi tendanciellement la règle juridique. A ce titre, la régulation par les données et les algorithmes mérite toute l'attention des pouvoirs publics, des juristes et des citoyens.

Pour en savoir plus :

Dossier : « Big Data entre régulation et architecture », *Statistique et Société*, vol. 2, n°4, 2014.

Lyria B. Moses et J. Chan, « Using Big Data for Legal and Law Enforcement Decisions : Testing the New Tools », *University of New South Wales Law Journal*, vol. 37, n°2, pp. 643 – 678.

OECD, *Technologies for Better Tax Administration. A Practical Guide for Revenue Bodies*, Paris, OECD Publishing, 2016